Torregrosa, A., Correig, E., Colomer, M., Vilalta, A., & Herranz, M. (2026). Analyzing the Impact of a Mathematical Competency-Based Curriculum: The Innovamat Case. *Revista de Investigación Educativa, 44*. https://doi.org/10.6018/rie.631081

Analyzing the Impact of a Mathematical Competency-Based Curriculum: The Innovamat Case

Analizando el impacto de un programa curricular basado en el desarrollo de la competencia matemática: El caso de Innovamat

Alba Torregrosa*[1] , Eudald Correig**, Marc Colomer***, Albert Vilalta* and Marta Herranz***

*Department of Mathematics and Experimental Sciences Education. Autonomous University of Barcelona (Spain)
**Department of Engineering. Faculty of Engineering. Rovira i Virgili University (Spain)
***Department of Educational Research. Innovamat Education SL Company (Spain)

Abstract

This study examines the impact of Innovamat: a curricular teaching and learning programme that aims to foster students' mathematical competence. To do so, this research considers the acquisition of mathematical competence in two different samples of 4th grade Mexican primary school students. Both samples belong to private schools with similar socio-economic backgrounds, their HDI (Human Development Index) being between 0.7 and 0.9. The first sample, called the intervention group, consists of 28 private schools (912 pupils) that have used Innovamat for a year. The second sample, called the control group, consists of 27 schools (1,111 pupils) that use textbooks from publishers widely known in Mexico. Our results indicate a correlation between the use of this programme—Innovamat—and better performance on an adapted version of the TIMSS test (Trends in International Mathematics and Science Study), both in terms of content and in cognitive domains that characterise mathematical competence. However, it is not possible to conclude that the implementation of this programme improves student mathematical competence, since this study does not allow us to establish its causal effect.

Keywords: mathematics; problem solving; instructional materials; evaluation; primary education.

Resumen

En el presente estudio analizamos el impacto de Innovamat: un programa curricular de enseñanza y aprendizaje que pretende desarrollar la competencia matemática del alumnado. Para ello, tomamos como objetivo medir el grado de adquisición de la competencia matemática de dos muestras de estudiantes mexicanos de 4º grado de educación primaria. Ambas muestras pertenecen a colegios privados similares a nivel socioeconómico, con un índice IDH (Índice de Desarrollo Humano) entre 0.7

[1] **Correspondence:** Alba Torregrosa,alba.torregrosa@uab.cat, Department of Mathematics and Experimental Sciences Education, Faculty of Education, Building G5, Plaça del Coneixement, 08193 Bellaterra, Barcelona (Spain).

y 0.9. La primera muestra, llamada grupo intervención, está conformada por 28 colegios (912 estudiantes) que usan el programa de Innovamat desde hace un año y la segunda muestra, llamada grupo control está conformada por 27 colegios (1,111 estudiantes) que usan editoriales de libros de texto muy extendidas en el territorio mexicano. Los resultados apuntan a una asociación entre usar el programa de Innovamat y un mejor desempeño en una adaptación de la prueba TIMSS (Trends in International Mathematics and Science Study) tanto en contenidos, como en dominios cognitivos que caracterizan la competencia matemática. A pesar de ello, no es posible afirmar que la implementación de este programa mejore la competencia matemática del alumnado, dado que el presente estudio no permite establecer su efecto causal.

Palabras clave: matemáticas; resolución de problemas; material didáctico; evaluación de la educación; educación primaria.

Introduction and objectives

One of the disciplines that has most rethought its purpose over the years is mathematics education. The question "what is doing mathematics" has dominated much of the research in educational literature for more than 50 years. Today, there is broad agreement among education experts on what learning mathematics entails (Silver et al., 1990), but we still find a dichotomy between what the scientific literature says doing mathematics is and what happens in the classroom.

Understanding that mathematics goes beyond the development of content is not a new or experimental trend. As early as the 1940s, Pólya (1945) mentioned problem solving as the central axis of mathematical activity. In the 1970s, Freudenthal (1973) mentioned that mathematical activity should promote the development of mathematical processes linked to mathematical competence. Currently, authors such as Alsina (2012, 2019), Casey and Sturgis (2018) and Liljedahl (2020) emphasise the imperative need to enhance students' mathematical competence in order to train students and future citizens to be more effective in dealing with the real problems posed by everyday life, beyond strictly academic ones. This competency-based view of mathematics has influenced much of the educational curriculum and organisations specialising in the teaching and learning of mathematics (NCTM, 2000; NGACBP and CCSSO, 2010; OECD, 2017), which understand that learning mathematics requires the development of key mathematical processes, such as problem solving, communication and representation, connections, and reasoning.

But are we really developing mathematical competence in the classroom? Recent studies (Martins and Martinho, 2024) show that there are still a large number of mathematics curricula that prioritise algorithmic application exercises over problem solving.

In the specific case of Mexico, the results of the PISA (*Programme for International Student Assessment*) tests have been on a downward trend for more than 10 years (Schleicher, 2019). The latest results from the 2022 edition of PISA place the country at a low level of performance in mathematical competence, scoring 395 points, far from the OECD average of 472 points. In addition, the school materials that support the teaching-learning process in Mexico, mostly conventional textbooks, often prioritise mechanical exercises that require the direct application of mathematical algorithms, rather than rich problems that allow for a greater variety of approaches and solutions (Valencia Álvarez and Valenzuela González, 2017).

Given this context, it is important to understand the effect of implementing curriculum programmes in Mexican classrooms that focus on the development of mathematical competence. In this study, we analyse the case of Innovamat, a mathematics teaching and learning curriculum programme that was first introduced in Mexico in the 2022-2023 school year and which defines itself as a programme to promote mathematical competence among students (Vilalta, 2021).

Recently, the authors of Innovamat published a white paper on mathematics entitled "Learning mathematics: Theoretical foundations of the Innovamat proposal" (Innovamat, 2024). In this white paper, the authors outline the fundamental pillars that underpin the programme and clarify the main axes for the development of mathematical competence. Despite this, the results of implementing Innovamat in real classrooms have not yet been published in research settings.

The aim of this study is to provide preliminary evidence on the relationship between using the Innovamat programme in the classroom and student performance on a competency-based mathematics test. In this regard, we aim to answer the question: Are students who use Innovamat better able to deal with situations that require skills related to mathematical competence, compared to students who work on mathematics with more conventional textbooks?

To this end, this study focuses on examining the performance of two different samples of 4th grade primary school students on an adaptation of the TIMSS (Trends in International Mathematics and Science Study) standardised mathematics test. The first sample, called the intervention group, consists of 28 private schools (912 students) that have been using the Innovamat programme for a year, and the second sample, called the control group, consists of 27 private schools (1,111 students) that use textbooks widely used in Mexico. The selection and adaptation of the TIMSS test as a measure is due to two main reasons. First, it is the most relevant mathematics test in primary education at the international level. Second, its theoretical framework includes a competency- pproach to mathematics, with activities that involve rich contexts on which it is important to reason in order to arrive at a solution (Barroso et al., 2021; Suárez-Pellicioni et al., 2016; Teig et al., 2022). Although the study design does not allow us to identify the causal effect that the Innovamat programme has on students' mathematical performance, it does allow us to obtain preliminary evidence on the relationship between the use of the programme and the results in an adaptation of a mathematics competency test.

The Innovamat programme

Innovamat is a mathematics curriculum programme for early childhood, primary and secondary education. Its objective is to design teaching resources based on research and teaching experiences that enable the development of mathematical competence (Vilalta, 2021). The programme defines mathematical competence in the words of Niss and Højgaard (2019): "mathematical competence is a person's conscious readiness to act appropriately in response to a specific type of mathematical challenge in given situations" (p. 6). The development of this competence involves developing four mathematical processes: problem solving, reasoning and proof, connections, and communication and representation (Santos-Trigo, 2024).

In this sense, and unlike conventional textbooks, the Innovamat programme makes a clear commitment to the development of mathematical processes and introduces mathematical content as an essential basis for working on these processes (Vilalta, 2021). Based on this, the programme proposes to develop mathematical competence through the combination of mathematical processes and content blocks (see Figure 1).

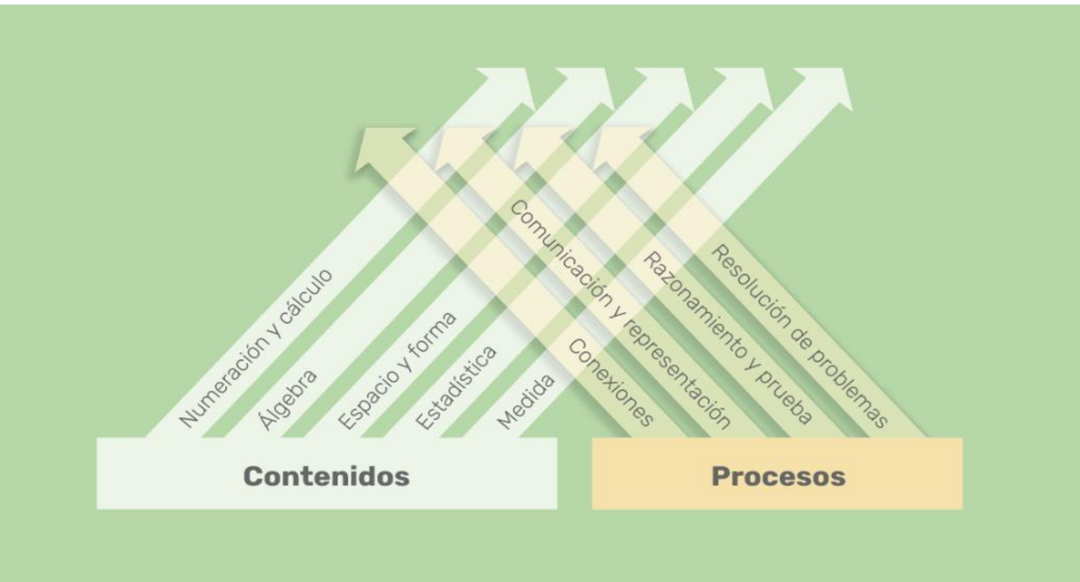


Figure 1. *Representation of the combination of content and processes*
Note. Innovamat (2024).

Innovamat is delivered to schools in the form of teaching guides for teachers, accompanied by individual record books for students, manipulative materials, digital resources, and training. Each of the above materials will be specified and justified in detail below, following the guidelines set out by Vilalta (2021) and the authors of Innovamat (2024).

The teaching guides for teachers contain an explanation of the activities to be carried out during a class session, as well as the use of manipulative materials and digital resources necessary to carry out these activities.

The activities contained in the Innovamat teaching guides follow the socio-constructivist framework based on the idea that students weave their learning through interaction with their environment and with others (Arcavi, 1999; Stephan and Akyuz, 2022). By providing teaching guides aimed solely at teachers, they take on a key role as mentors (Banchi and Bell, 2008), facilitating the discovery of knowledge by students ( ) and, in this process, developing their ability to solve problems, reason, communicate mathematical concepts and make connections between concepts or with their environment. The activities, as well as the teaching role, promote guided discovery, based on the idea of creating opportunities to develop mathematical processes and students' self-confidence. Other educational programmes that promote guided discovery, such as JUMP Math, have been shown to have a positive impact on the mathematical performance of primary school students (Solomon et al., 2019).

The teaching sequences that organise the activities in the teacher's guides follow the CRA (Concrete, Representative and Abstract) model (Laski et al., 2015; Shuxratovna, 2024). In an initial period, the content is discovered through manipulation and experimentation, which gives way in a second phase to the pictorial representation of the materials used. This path ends in abstraction, which pursues the goal of discovering increasingly symbolic algorithms or strategies. As an example, Figure 2 shows the teaching sequence proposed by Innovamat to move from manipulation to the conventional algorithm of addition. The CRA model has been shown to have a positive effect on students' conceptual understanding of mathematics and their mathematical representation (Purwadi et al., 2019), particularly for students with learning difficulties (Bouck et al., 2018). The CRA model is also a fundamental aspect of the curriculum in Singapore, one of the countries that has achieved the best results in PISA tests over the last decade (Leong et al., 2015).

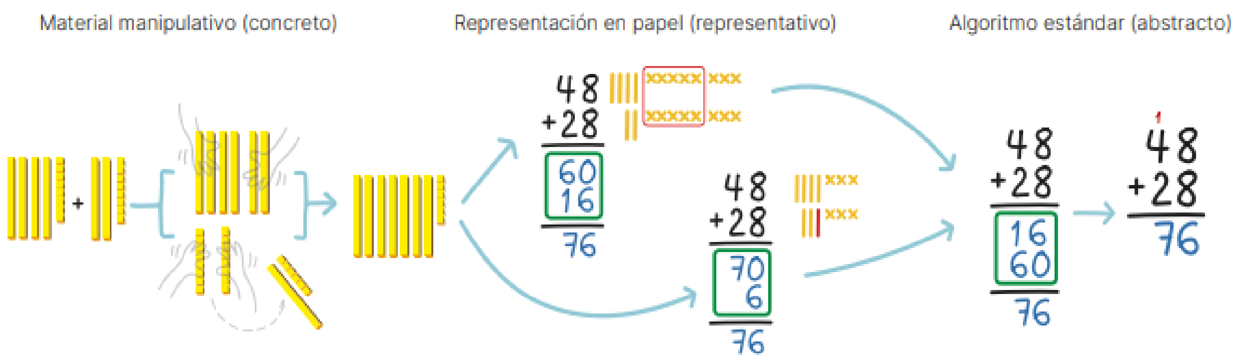


Figure 2. *Example of the construction of the sum decomposition strategy in Innovamat, using the CRA model*
Note. Innovamat (2024)

In addition to objective explanations of the activities based on the CRA model, the teaching guides for teachers include possible questions that teachers can ask students (with possible answers), adaptations to increase or decrease the difficulty of the activities, and training tips. These elements promote Booth and Ainscow's (2002) three Ps, which emphasise Presence (physical inclusion of students), Participation (social and emotional) and Progress (academic inclusion).

As an example, Figure 3 shows a teaching guide for teachers on a Numbering and Calculation session for 3rd grade primary school, which is structured in three different stages. The initial stage, or Warm-up, in which students are asked to do an activity to refresh their prior knowledge or anticipate the knowledge they will develop during the session (Skemp, 1976). The central part, or Let's talk, in which the aim is for students to construct mathematical knowledge in groups, based on a question or challenge posed by the teacher (Piggott, 2011): using manipulatives (interlocking cubes, segments, base 10 blocks, geoboards, etc.) and/or digital resources (videos, projectable presentations, songs, etc.). And the final moment or Let's Record, in which students are asked to tackle individual mathematical problems in their record notebooks. This session structure aims to help students build knowledge collectively without forgetting the CRA model (Laski et al., 2015), which is necessary for consolidating mathematical learning individually (Schoenfeld, 2014). The guides suggest spending 35 to 45 minutes on the group development of Let's warm up and Let's talk. The last 10 to 15 minutes of the session will be devoted to Let's Record in the students' individual notebooks.

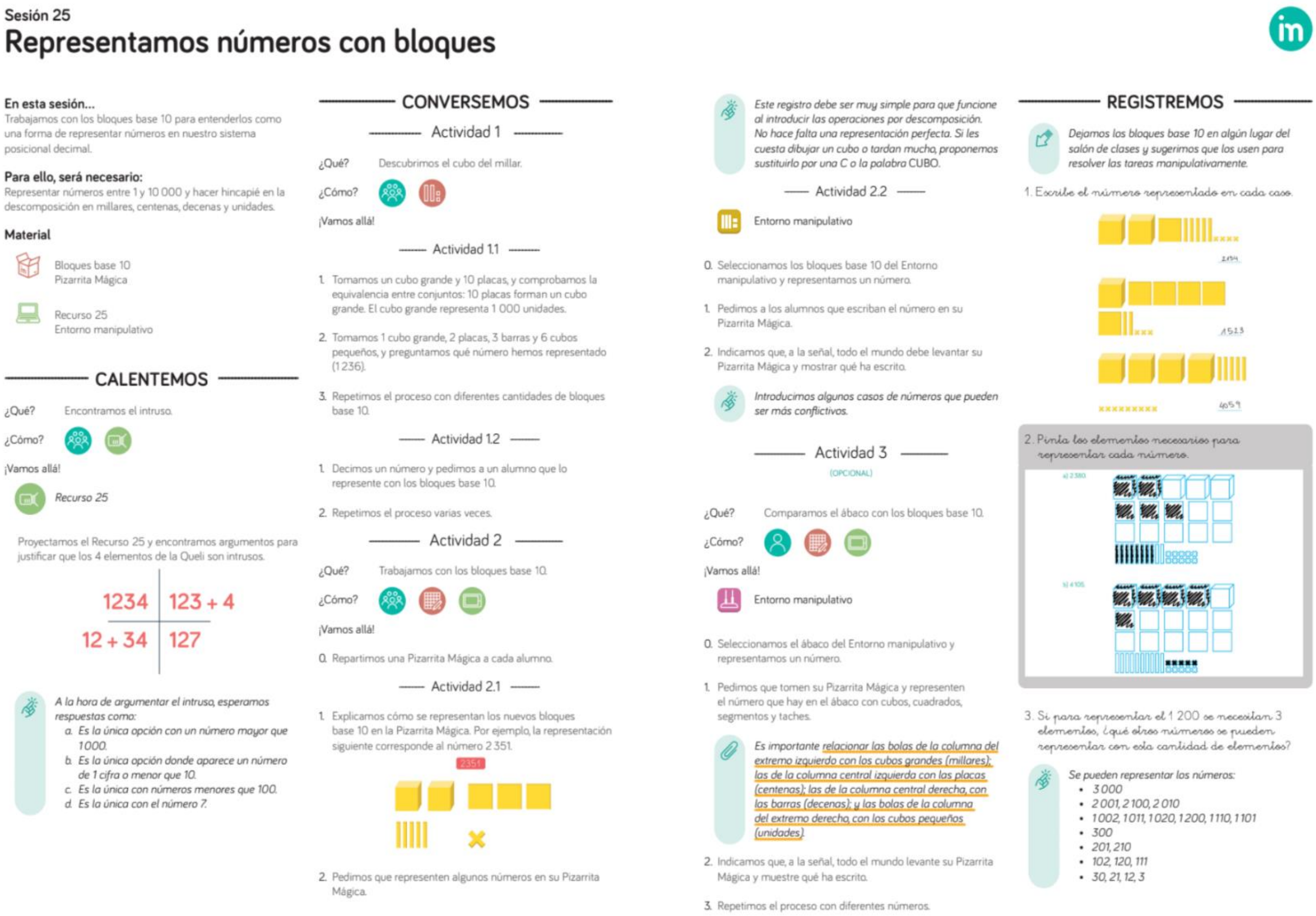


Figure 3. *Session 25 of the teaching guide for 3rd grade primary school*
Note. Innovamat (2024).

Each school year contains approximately 85 class sessions. As detailed above, each session has its own teaching guide for teachers and its own individual record book for students (in addition to the digital and manipulative resources that the school receives and that allow each session to be developed). The programme includes approximately three sessions per week plus one additional weekly session dedicated to systematic digital practice. This systematic digital practice takes place in a gamified and self-adaptive environment in which students must complete systematic activities or exercises (Mayer, 2002) to build their own city. This weekly hour dedicated to systematic digital practice aims to consolidate conceptual and procedural learning, as well as to develop fluency, efficiency and flexibility in manipulating and operating with numbers (Bay-Williams and SanGiovanni, 2022). In this sense, systematic digital practice aims to contribute to a deeper understanding of mathematical fundamentals, build confidence and help develop the ability to estimate and verify the reasonableness of results (Schoenfeld, 2014).

The order in which class sessions are presented within a single course, as well as the relationship between sessions from different courses, is based on the construction of spiral sequences (Bruner, 1977). Spiral learning sequences are those that take relational learning into account, seeking to build a solid conceptual structure that allows students to generate a variety of strategies and connections rather than understanding mathematics as separate concepts in unconnected blocks of content (Skemp, 1976). Unlike the programmes found in most textbooks, which are divided into thematic units that must be 'mastered' before moving on to the next, spiral trajectories review and deepen mathematical concepts over time, allowing for the diversity of student maturity to be addressed, thus reinforcing the connections between knowledge. Various studies (Howard, 2007; Kolomitro et al., 2017) indicate that spiral programmes or curricula promote a deeper understanding of mathematical concepts and processes . Such spiral sequences are also used by other programmes, such as the Singapore method (Thiyagu, 2013).

The type of tasks shown in Figure 4 are a good example of the interrelation between content or spiral sequencing. As we can see, in all four activities, students must write the sum of the two numbers in the previous row inside the rectangles and the subtraction of the two numbers in the previous row (left minus right) inside the hexagons. Any two numbers can go inside the circles (the larger number must go in the large circle). Once the diagrams are complete, students must explain the relationship they find between the numbers in the first and third rows.

As we can see in Figure 4, this activity is set in Year 3 (with natural numbers), Year 5 (with decimal numbers), Year 7 (with whole numbers) and Year 9 (with variables). Individual completion of the activity in the third year of secondary school allows students to arrive at the algebraic expressions that justify the conjectures about the relationship between the first and last rows of all the previous diagrams.

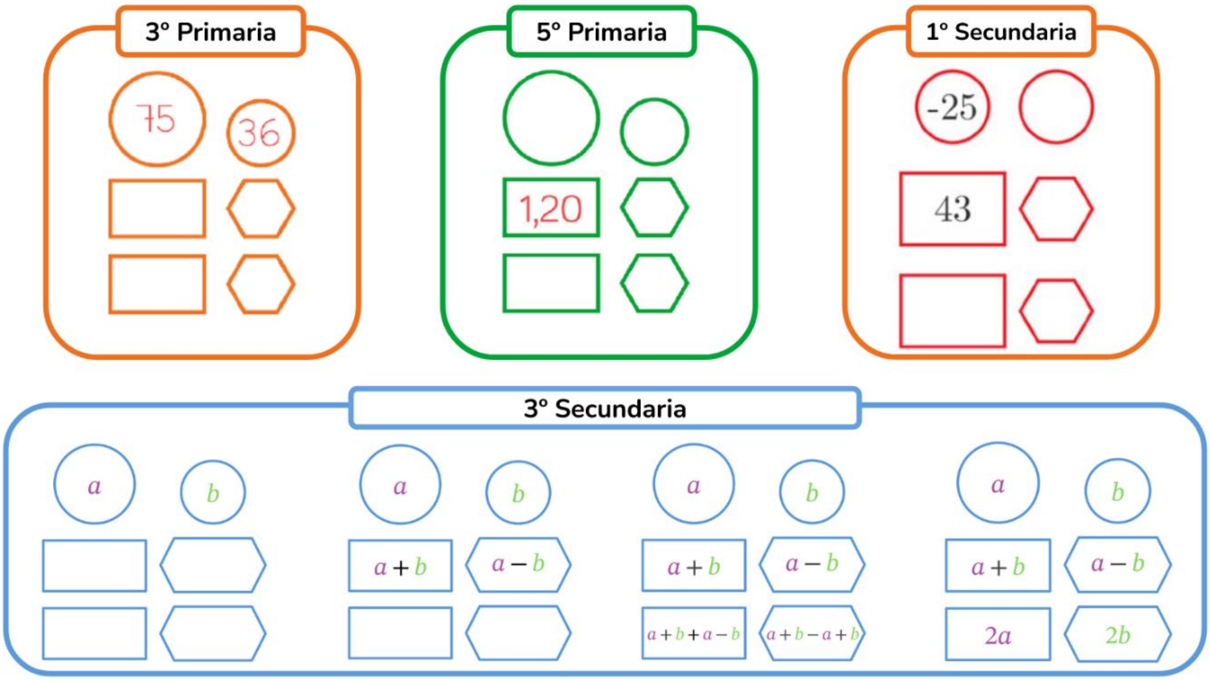


Figure 4. *Diagram activity taken from the individual records of four different courses*
Note. Innovamat (2024).

Having detailed the materials that Innovamat provides to educational centres and justified their design and implementation, we will now describe the teacher training that accompanies the programme.

Innovamat offers two types of teacher training: one in a large-scale online format for all teachers and another individualised face-to-face format for schools. Both training courses are based on the idea that the implementation of a mathematics teaching programme based on scientific evidence requires opportunities for teachers to gain an in-depth understanding of the programme's theoretical framework, as well as practical advice for its implementation and optimal development (Schulman, 1986). In this sense, Innovamat's teacher training proposal follows the MKT (Mathematical Knowledge for Teaching) framework, providing teachers with all the knowledge they need to transform the mathematics classroom (Ball et al., 2008).

Large-scale online training sessions are held three times a year, coinciding with the start of each term, and last one hour each. In these training sessions, the Innovamat teaching team provides an overview of each term of the course, focusing on understanding the spiral sequence with which the sessions that teachers will implement over the next three months have been constructed.

Individualised, face-to-face training sessions are held at each school and are conducted by an advisor who is a specialist in mathematics teaching. These training sessions are held every two months and last between one and two hours (depending on the school's needs). These training sessions have different purposes: to carry out classroom modelling or observations, to train teachers in teaching trajectories (Clements and Sarama, 2020) and specifics of the spiral sequence, to train teachers in the DUA framework (Meyer et al., 2014), etc. At the beginning of the school year, specialist advisors, together with the school management teams, agree on the topics to be covered in the training sessions and the specific timetable for them.

In parallel with the detailed training courses, the Innovamat teaching team holds an annual *webinar* open to families via the YouTube platform, in which the programme is presented in a clear and entertaining way. They also usually hold annual "thematic" *webinars* on topics of interest in the field of mathematics education (such as the explanation of the PISA 2022 results or the transition from primary to secondary school), which are open to the entire educational community.

Method

Participants

In this study, the participants are divided into two different samples. The first sample, called the intervention group, consists of 28 private schools (912 4th grade primary school students) that have switched from using widely used textbook publishers in Mexico to implementing the Innovamat programme for one school year. The intervention group includes more than 50% of schools that were using Innovamat materials in some region of Mexico at the time of data collection (28 schools out of 54 in total). In this sense, we consider the sample belonging to the intervention group to be representative given the total population of schools implementing the programme. The second sample, called the control group, consists of 27 private schools (1,111 fourth-grade primary school students) that continue to use the same textbook publishers as the intervention group previously used.

Given that the Innovamat programme is currently only implemented by private Mexican schools, all schools belonging to the intervention group share this characteristic. In order to maintain homogeneity in terms of school ownership, the control group is also entirely made up of privately owned schools.

To ensure that both groups were stable in terms of socioeconomic factors, the Human Development Index (HDI) developed by the United Nations Development Programme (UNDP) was used to compare the two groups. The HDI is a synthetic measure used to capture and compare the level of prosperity of nations. This index assesses fundamental aspects of human development through three main dimensions: life expectancy at birth, which reflects health and longevity; the level of education, measured by the average schooling rate and the expected schooling rate; and the standard of living, quantified by gross national income (GNI) per capita adjusted for purchasing power parity (UNDP, 2020). The HDI seeks to provide a more holistic framework than gross domestic product (GDP) per capita for understanding human well-being and progress (Stanton, 2007). The HDI is represented by a numerical scale ranging from 0 (lowest development) to 1 (maximum development). Both groups have an HDI between 0.7 and 0.9, so we can consider them homogeneous with respect to this indicator.

Both the schools in the intervention group and those in the control group are schools with one or two educational streams and approximately 25 students per classroom. All schools devote four hours per week to mathematics instruction. In terms of gender distribution, 54% of participants in the intervention group are female, while 46% of participants in the control group are female.

In line with the purpose of the study, which seeks to compare the degree of mathematical competence acquired by both groups, students with individualised educational support programmes were excluded from the study. In this regard, a total of 18 students from the intervention group and a total of 21 students from the control group did not participate in the study. This decision was made by the research team, in conjunction with the school management teams, given the mathematical and literacy difficulties presented by the adaptation of the TIMSS standardised test used in the study.

Based on the above indicators, we consider that both groups are homogeneous in terms of HDI, school size, students per classroom, gender, educational level and hours devoted to teaching and learning mathematics.

Instruments

The instrument used in this study to measure the degree of mathematical competence acquisition is an adaptation of the TIMSS test (Mullis et al., 2020). The members of the research team decided to use the TIMSS test because it has been widely used in research as it is one of the most validated standardised tests that assesses student performance in mathematical competence (Barroso et al., 2021; Suárez-Pellicioni et al., 2016; Teig et al., 2022). The scope of TIMSS is growing, and in its latest version in 2023, a total of 72 countries participated, including students in the 4th grade of primary education and the 2nd grade of secondary education. In this study, given the sample selection detailed above, we will focus on the test taken by 4th grade primary school students.

In the official TIMSS test for 4th grade primary education, students have about 36 minutes to answer 20 to 28 maths questions (depending on the set of items randomly assigned to each student) in which they must demonstrate their performance in the development of mathematical content and cognitive domains. The questions can take different formats, from selecting a closed answer to selecting multiple answers or writing open-ended answers.

In terms of the distribution of mathematical content, the TIMSS theoretical framework proposes tasks based on three main blocks:

- *Numbers:* 50% of the questions on the test are aimed at developing numbering and calculation content along with the development of algebraic thinking.
- *Geometry and measurement:* 35% of the questions on the test are aimed at developing content related to shape, space and measurement.
- *Data:* 15% of the questions on the test are aimed at developing content related to statistics and probability.

In terms of cognitive domains, they are also divided into three main sections:

- *Knowledge:* 40% of questions cover mathematical facts, concepts and procedures that students should know.
- *Application:* 40% of questions measure students' ability to apply facts, concepts and procedures in different situations.
- *Reasoning:* 20% of questions revolve around unfamiliar situations, rich contexts, and multi-level/multi-content problems.

Given that Mexico has not administered the TIMSS test since 2000 (Backhoff and Solano-Flores, 2003), the members of the research team for this study were forced to adapt the TIMSS test from another country. Given the curricular and linguistic similarities between Chile and Mexico, we used the open-ended questions from the 2019 Chilean TIMSS test as a reference, which included questions from the 4th grade for the years 2011 and 2015.

From the dozens of open-ended questions, we selected a subgroup of 26 questions, containing a total of 33 sub-questions. Firstly, for technical reasons, we selected tasks that involved only a closed-ended response selection process (single or multiple choice) or tasks that allowed simple number entries to be solved, i.e., tasks that did not involve graphical representations or handwritten reasoning or other mathematical representations. This criterion was followed in order to adapt the test to the Typeform digital platform, which did not allow for the completion of extensive text fields. Secondly, the selection of questions was made in such a way that the distribution of questions for each content block and cognitive domain respected the percentages indicated above. In cases where more than one question could be included, the selection was made at random, as is done by the organisers of the original TIMSS test. A linguistic expert reviewed and refined the 33 sub-questions resulting from the adapted test in order to adjust some terms widely used in Chile but unknown in Mexico (such as the word *sticker* to refer to a stamp or adhesive). Table 1 shows the characteristics of each selected question and sub-question.

Regarding the test administration time, we decided to extend the time from 36 to 60 minutes. This decision was made for two main reasons. The first was to ensure that both the control group and the intervention group students could complete the entire test. This would allow us to compare the scores of both groups on all questions and sub-questions in the adapted TIMSS test. The second reason was to reduce the math anxiety that students may experience when faced with graded tests with a very limited time limit (Barroso et al., 2021; Castillo-Sánchez et al., 2020; López-Chao et al., 2020; Suárez-Pellicioni et al., 2016).

Table 1

Content block (C) and cognitive domain (D) of each question and sub-question (P) in the adaptation of the test used in this study.

P	C	D
1A	Numbers	Application
1B	Numbers	Reasoning
1C	Numbers	Reasoning
1D	Numbers	Reasoning
2	Numbers	Reasoning
3	Numbers	Reasoning
4	Numbers	Application
5	Numbers	Application
6	Numbers	Knowledge
7	Numbers	Application
8	Numbers	Knowledge
9	Numbers	Application
10	Numbers	Knowledge
11A	Numbers	Application
11B	Numbers	Application
12	Numbers	Application
13	Measurement and geometry	Knowledge
14	Measurement and geometry	Knowledge
15	Measurement and geometry	Application
16	Measurement and geometry	Application
17	Measurement and geometry	Application
18	Measurement and geometry	Application
19	Measurement and geometry	Knowledge
20	Measurement and geometry	Knowledge
21A	Measurement and geometry	Knowledge
21B	Measurement and geometry	Knowledge
21C	Measurement and geometry	Knowledge

21D	Measurement and geometry	Knowledge
22	Measurement and geometry	Application
23	Measurement and geometry	Application
24	Data	Knowledge
25	Data	Reasoning
26	Data	Knowledge

Procedure

The TIMSS test was made available to schools via a link to the Typeform digital platform for administering questionnaires and tests. This platform has a very intuitive, simple and user-friendly interface for students aged 9 to 10. Teachers had 14 calendar days to administer the test in their respective computer rooms. Before the test was sent, schools were informed about the purpose of the study, the anonymity clauses and the implementation protocol. Under no circumstances were teachers allowed to clarify the questions or provide support to students during the test. The test had to be taken by the students participating in the study completely independently and individually.

Once the 60 minutes of the test had elapsed, the students had to finish, even if they had not completed all the questions. The research team automatically received the data school by school and student by student (both the answers marked and the total time taken to complete the test). This data was received through the same Typeform digital platform from the administrator user. The research team verified that the test duration did not exceed 60 minutes in any case.

Because the original TIMSS test validates performance in content and cognitive domains that students should have acquired by the end of Year 4 of primary education, the adaptation of the TIMSS test used in this study was administered in a single edition in May 2023. Below are some examples of questions taken from the adaptation of the TIMSS test for Year 4 that address the development of different mathematical content and domains.

Table 2

Excerpt of questions from the "Knowledge" domain for each content block.

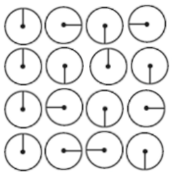
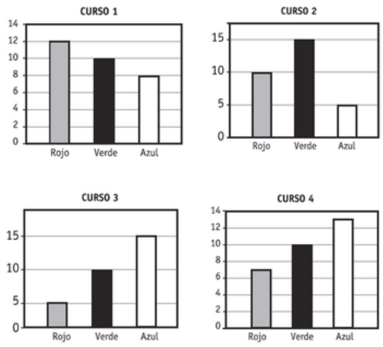
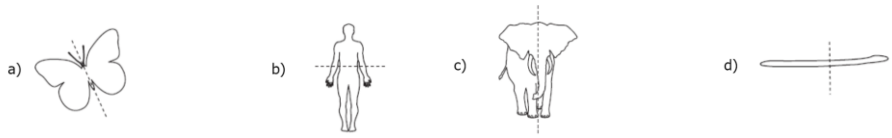
Numbers	Which gives the result closest to 9 x 22? a) 5 x 20 b) 5 x 25 c) 10 x 20 d) 10 x 25
Measurement and geometry	The rule for a sequence says, "Rotate the shape ¼ turn at a time, clockwise." What will the sequence look like? a) b) c) d) 
Data	Andrew has conducted a survey on the favourite colour of students in four year groups.  In which year are there fewer students who prefer the colour blue? a) Year 1 b) Class 2 c) Class 3 d) Class 4

Table 3

Excerpt of questions from the "Application" domain for each content block.

Numbers	Joana had 12 apples. She ate some and had 9 left. Which numerical expression describes what happened? a) 12 + 9 = __ b) 9 = 12 + __ c) 12 - __ = 9 d) 9 - __ = 12
Measurement and geometry	In which of these drawings is the dotted line an axis of symmetry? 

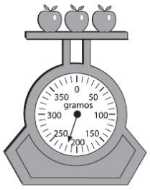
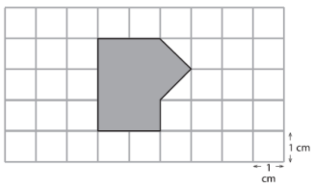
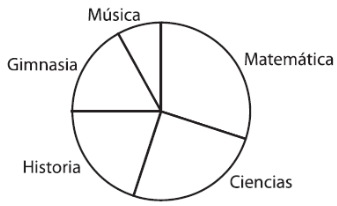
Data	 How many grams do the apples weigh? a) 200 b) 202 c) 210 d) 220

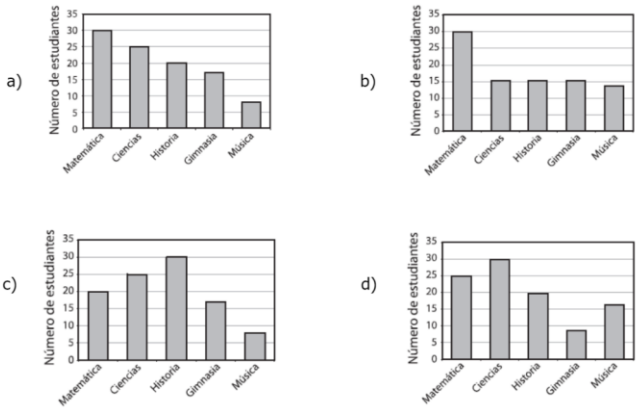
Table 4

Excerpt of questions from the "Reasoning" domain for each content block.

Numbers	Maria left El Almendro and cycled at the same speed for 2 hours. She reached this sign.  Maria continues cycling at the same speed to Buenaventura. How many hours will it take her to cycle from the sign to Buenaventura? a) 1 ½ hours b) 2 hours c) 3 hours d) 3 ½ hours
Measurement and geometry	The squares in the grid below are 1 cm by 1 cm.  What is the shaded area, in square centimetres?
Data	Mr. Rodriguez asked the students at his school about their favourite subject. The following pie chart shows how many students like each of the 5 subjects.



Which bar graph shows the same information as the pie chart?



Data analysis

The test was marked automatically using the Typeform platform, and the results for each student in each content block and cognitive domain were analysed by the research team for this study. The results are given as mean and standard deviation. For hypothesis testing, a Welch (1947) t-test is used for each content and domain, and Cohen's (1988) d is used to calculate effect sizes. The analyses were performed using R statistical software version 4.2.

Results

If we analyse the Human Development Index (HDI) of the 28 private schools that apply the Innovamat programme (intervention group) and the 27 private schools that use textbooks (control group), we find that the mean of the former (with a mean of 0.81 and a standard deviation of 0.03; $p = 0.57$), is not significantly different from the mean of the latter (with a mean of 0.80 and a standard deviation of 0.04). Thus, we can consider the two groups of schools to be equivalent in terms of HDI and other characteristics noted in the previous section. This allows us to compare the two groups directly.

With regard to the results, in the evaluation of the content blocks of the TIMSS test adaptation (Figure 5 and Table 5), the students in the intervention group—who use the Innovamat programme—outperformed the students in the control group—who use textbooks—in all the content blocks evaluated. The differences are statistically significant ($p < 0.001$) and have moderate effect sizes (Cohen's "d" > 0.2) in all three content blocks.

Figure 5. *Results in the different content blocks of the TIMSS test adaptation for students in the intervention group (Innovamat) versus students in the control group (textbook).*

Note: The graphs show the mean +/- standard error of the results for each content block. The value for each block represents the mean percentage of correct answers for the students in each group.

Table 5

Averages and standard deviations (in parentheses) for the results in the different content blocks of the TIMSS test adaptation for students in the intervention group (Innovamat) versus students in the control group (textbook).

Content block	**Intervention group. Innovamat (n = 912)**	**Control group. Textbook (n = 1111)**	**p-value**
Data	2.14 (0.86)	1.91 (0.89)	< 0.001
Measurement and geometry	9.06 (2.83)	8.46 (2.97)	< 0.001
Numbers	9.34 (3.42)	8.38 (3.67)	< 0.001

Note. The values shown are based on a total of 3 questions for "Data", 14 questions for "Measurement and Geometry" and 16 questions for "Numbers". The table shows the p-value of a T-test for comparing means, as well as the effect size calculated using Cohen's "d".

If we analyse the results in terms of cognitive domains (Figure 6 and Table 6), the results are similar. We found that the students in the intervention group outperformed the students in the control group in the three cognitive domains in a statistically significant way ($p < 0.001$) and also with moderate effect sizes (Cohen's "d" > 0.2).

Figure 6. *Results in the different cognitive domains of the TIMSS test adaptation for students in the intervention group (Innovamat) versus students in the control group (textbook).*

Note: The graphs show the mean +/- standard error of the results for each content block. The value for each cognitive domain represents the mean percentage of correct answers for the students in each group.

Table 6

Averages and standard deviations (in parentheses) for the results in the different cognitive domains of the TIMSS test adaptation for students in the intervention group (Innovamat) versus students in the control group (textbook).

Cognitive domain	Intervention group. Innovamat (n = 912)	Control group. Textbook (n = 1111)	p-value
Application	8.45 (2.97)	7.68 (3.22)	< 0.001
Knowledge	8.80 (2.55)	8.07 (2.72)	< 0.001
Reasoning	3.16 (1.60)	2.79 (1.65)	< 0.001

Note: The values shown are based on a total of 14 questions for "Application", 13 questions for "Knowledge", and 6 questions for "Reasoning". The table shows the p-value of a T-test for comparing means, as well as the effect size calculated using Cohen's "d".

Discussion and conclusions

The results of this study suggest that the implementation of the Innovamat programme is related to better student performance on the adapted TIMSS (Trends in International Mathematics and Science Study) standardised test. Schools that followed the Innovamat programme for one year (intervention group) scored significantly better on the adapted TIMSS test than schools that followed a textbook-based teaching model (control group).

Statistical analysis found Cohen's effect sizes ranging from 0.21 to 0.28 for the different domains assessed (Tables 5 and 6). In the existing literature, the effect sizes of educational interventions tend to be modest. For example, Hattie (2008) in his meta-analysis synthesis points out that the average effect size of educational interventions is around 0.40. However, these effects accumulate over several years of instruction (Hattie, 2008).

Over the course of a school year, effect sizes are typically smaller, and even an effect of 0.20 is considered large and, in many cases, uncommon (Kraft, 2020). In this regard, the effect sizes in the present study slightly exceed the threshold of 0.20 in all content blocks and cognitive domains, although these results should be interpreted with caution.

Firstly, there are a multitude of factors that can influence the data obtained. As Tashtoush et al. (2022) point out, there are multiple issues that can alter the results obtained by both samples in the TIMSS test adaptation: such as the general methodological approach of the school, the leadership of the management team, the quality of teacher training received over time, the fact that the test was conducted in a digital format that only takes into account the correctness of the answer, the extension of the test application time, or other socio-economic factors that are not included in the HDI (Human Development Index). In this regard, it is not possible to separate the influence that the Innovamat programme may have had on students' mathematical learning from other relevant factors. Although implementing the Innovamat programme seems to be related to better performance in the TIMSS test adaptation, it is possible that factors other than Innovamat may have influenced the result.

Secondly, we would like to highlight the limitation of not having a pre-test carried out at the beginning of the school year. In this regard, the study shows preliminary evidence that the use of Innovamat materials is related to better performance on an adaptation of the TIMSS test, by comparing two samples of schools with similar socioeconomic levels. In other words, under similar socioeconomic conditions, schools that followed the Innovamat programme obtained better results at the end of the school year in a test that assesses the acquisition of certain curricular content, as well as the ability to apply knowledge and reason in unfamiliar situations. The design of this study allows us to relate the use of a certain programme to test results, but it does not allow us to evaluate what changes have occurred in the students due to the programme analysed. To overcome this limitation, we believe that future research could use an experimental design with pre-tests and post-tests to better understand what changes are generated by following the Innovamat programme. Specifically, to better understand the causal impact of the programme, a randomised controlled trial (RCT) design could be used. Although this design is difficult to implement and requires a large amount of financial and logistical resources, it is the gold standard for identifying the causal effect of implementing an educational programme (Connolly et al., 2018). Based on the results of this study, we would expect the Innovamat programme to have a positive causal impact on student learning.

Thirdly, we consider it relevant to mention the short duration of the study, with only one year of programme implementation. Although the results seem promising given the effect size obtained (Kraft, 2020), it would be interesting to analyse how this effect evolves over a longer period of exposure. Slavin et al. (2011) point out that methodological changes in teaching and learning often take several years to reflect statistically significant improvements. Since the Innovamat programme includes spiral learning trajectories for the entire educational stage, from 3 to 16 years of age, future studies could compare students' mathematical performance according to the number of years they have been following the programme.

Finally, we would also like to highlight that the study was carried out in a specific context: private schools in Mexico. Opening the way for future research, it would be interesting to analyse whether the results of this study can be generalised to other educational contexts, such as public schools. Our conjecture is that the implementation of competency-based models could be equally effective in diverse educational environments. Studies such as that by Stigler and Hiebert (1999) show that different educational systems can benefit significantly from innovative teaching approaches, which supports the possibility of successfully adapting the Innovamat programme to different educational contexts.

Returning to the study's research question and given the results presented above, we consider that the use of Innovamat materials is related to better performance on the TIMSS test, and this effect is not specific to any particular content block or cognitive domain, but rather generalised. Although these results are preliminary, as we have argued previously, they raise a clear question: what characteristics of the Innovamat programme could be related to the good results obtained?

We hypothesise that there are two key factors in the Innovamat programme that directly address the development of mathematical competence. On the one hand, the spiral design (Howard, 2007; Kolomitro et al., 2017) of the teaching guides, the record books and the digital application of the proposal. We believe that this design allows students to build a solid conceptual structure, generating a wide variety of strategies and connections (Skemp, 1976) similar to those required by the adapted TIMSS test.

On the other hand, the fact that teacher training is provided within the framework of MKT (*Mathematical Knowledge for Teaching*) (Ball et al., 2008). In this regard, we believe that the Innovamat programme is correct in considering that it is not possible to develop students' mathematical competence if teachers are not mathematically competent. To this end, we believe that both online and face-to-face training, as well as the design of the teaching guides themselves, could have a positive effect on the teaching and learning of mathematics.

Returning to the ideas put forward by Alsina (2012, 2019), Casey and Sturgis (2018) and Liljedahl (2020), as well as those highlighted by organisations such as the OECD (2017), there is a growing international demand for educational approaches that foster critical and problem-solving skills, which are central elements in competency-based teaching. Furthermore, as reflected by Bakker et al. (2023), it is important that research in mathematics education in the coming years focuses on competency-based teaching approaches, the objectives of mathematics education, and equity, diversity and inclusion. Therefore, we consider it interesting to highlight the relevance of the results obtained in this study, especially given Mexico's performance in the PISA (*Programme for International Student Assessment)* mathematics tests. As Darling-Hammond and Adamson (2014) suggest, incorporating innovative assessment and teaching approaches into education policies may be crucial for the advancement of more effective and relevant education systems in the 21st century.

In conclusion, we believe that we have achieved the objective of the study by providing preliminary evidence on the relationship between using the Innovamat programme in the classroom and student performance on a competency-based mathematics test. In this regard, and returning to the study's research question, we believe that students who have been using the Innovamat programme for a year show a greater ability to deal with situations that require skills related to mathematical competence, compared to students who work on mathematics with more conventional textbooks. The implications of this study encourage a paradigm shift towards teaching based on the development of mathematical competence, which not only aligns instruction with the demands of international educational standards, but also promotes deeper and more applicable mathematical learning. We note that these results are relevant to the educational community, as they inform the potential benefits of competency-based programmes and bring the didactic foundations of these programmes closer to the reality of the classroom.

References

Alsina, Á. P. (2012). Beyond content: mathematical processes in early childhood education. *Edma 0-6: Mathematics Education in Childhood, 1*(1), 1-14. https://doi.org/10.24197/edmain.1.2012.1-14

Alsina, Á. P. (2019). *Teaching itineraries for mathematics (6-12 years).* Editorial Graó.

Arcavi, A. (1999). ...And in mathematics, what do we build when we teach? *Numbers. Journal of Mathematics Education, 38,* 39-56. Retrieved from: https://dialnet.unirioja.es/servlet/articulo?codigo=2343631

Backhoff, E. and Solano-Flores, G. (2003). *Third International Study of Mathematics and Natural Sciences (TIMSS): Mexico's results in 1995 and 2000. No. 4*. National Institute for the Evaluation of Education.

Bakker, A., Cai, J., and Zenger, L. (2023). Future topics of research in mathematics education: an international survey before and during the pandemic. *Mathematics Education, 35*(2), 9–46. https://doi.org/10.24844/EM3502.01

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching. *Journal of Teacher Education, 59*(5), 389–407. https://doi.org/10.1177/0022487108324554

Banchi, H., & Bell, R. (2008). The many levels of inquiry. *Science and children, 46*(2), 26. https://hal.science/hal-00692073v1

Barroso, C., Ganley, C. M., McGraw, A. L., Geer, E. A., Hart, S. A., & Daucourt, M. C. (2021). A Meta-Analysis of the Relation Between Math Anxiety and Math Achievement. *Psychological Bulletin, 147*(2), 134-168. https://doi.org/10.1037/bul0000307

Bay-Williams, J. M. and SanGiovanni, J. J. (23 April 2022). *Accessing fluency through routine and opportunity* [presentation]. NCTM Annual Meeting 2022, Los Angeles, CA.

Booth, T. and Ainscow, M. (2002). *Index for Inclusion. Developing learning and participation in schools. Centre for Studies on Inclusive Learning (CSIE)*.

Bouck, E. C., Satsangi, R., and Park, J. (2018). The Concrete–Representational–Abstract Approach for Students With Learning Disabilities: An Evidence-Based Practice Synthesis. *Remedial and Special Education, 39*(4), 211–228. https://doi.org/10.1177/0741932517721712

Bruner, J. S. (1977). *The process of education.* Harvard University Press.

Casey, K., & Sturgis, C. (2018). *Levers and Logic Models: A Framework to Guide Research and Design of High-Quality Competency-Based Education Systems.* iNACOL. Retrieved from: https://bit.ly/3w0QeE9

Castillo-Sánchez, M., Gamboa-Araya, R., & Hidalgo-Mora, R. (2020). Factors influencing student dropout and failure in a university mathematics course. *Uniciencia, 34*(1), 219-245. https://doi.org/10.15359/ru.34-1.13

Clements, D. H. and Sarama, J. (2020). *Learning and teaching early math: The learning trajectories approach.* Routledge.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed). L. Erlbaum Associates. https://doi.org/10.4324/9780203771587

Connolly, P., Keenan, C., & Urbanska, K. (2018). The trials of evidence-based practice in education: a systematic review of randomised controlled trials in education research 1980–2016. *Educational Research, 60*(3), 276–291. https://doi.org/10.1080/00131881.2018.1493353

Darling-Hammond, L. and Adamson, F. (2014). *Beyond the Bubble Test: How Performance Assessments Support 21st Century Learning*. John Wiley & Sons. https://doi.org/10.1002/9781119210863

Freudenthal, H. (1973). *Mathematics as an educational task*. Kluwer Academic Publishers.

Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement.* Routledge. https://doi.org/10.4324/9780203887332

Howard, J. (2007). *Curriculum development.* Centre for the Advancement of Teaching and Learning. Elon University.

Innovamat Education S.L. (3 December 2024). *White paper: Learning mathematics. Theoretical foundations of Innovamat's proposal.* https://www.innovamat.com/wp-content/uploads/2025/04/Libro-blanco-aprendizaje-matematicas-Innovamat.pdf

Kolomitro, K., Inglese, J., & Idzikowski, M. (2017). *Curriculum Design Handbook.* Centre for Teaching and Learning. Queen's University. http://www.dgma.donetsk.ua/docs/kafedry/avp/metod/Kolomitro%20-%20Curriculum%20Design%20Handbook%20-%202017.pdf

Kraft, M. A. (2020). Interpreting Effect Sizes of Education Interventions. *Educational Researcher*, *49*(4), 241–253. https://doi.org/10.3102/0013189X20912798

Laski, E. V., Jordan, J. R., Daoust, C. J., & Murray, A. (2015). What Makes Mathematics Manipulatives Effective? Lessons From Cognitive Science and Montessori Education. *SAGE Open*, *5*(2), 1–8. https://doi.org/10.1177/2158244015589588

Leong, Y. H., Ho, W. K., & Cheng, L. P. (2015). Concrete-Pictorial-Abstract: Surveying its origins and charting its future. *The Mathematics Educator, 16*(1), 1-18. https://hdl.handle.net/10497/18889

López-Chao, V., Mato-Vázquez, D., & Chao-Fernández, R. (2020). Confirmatory analysis of the factorial structure of anxiety towards mathematics. *Journal of Educational Research, 38*(1), 221–237. https://doi.org/10.6018/rie.359991

Liljedahl, P. (2020). *Building Thinking Classrooms.* Corwin.

Martins, L. G. and Martinho, M. H. (2024). Types of tasks in mathematics textbooks: a study of Portuguese textbooks for 10th and 11th grades. *Mathematics Education, 36*(1), 66-91. https://doi.org/10.24844/EM3601.03

Mayer, R. E. (2002). Rote Versus Meaningful Learning. *Theory Into Practice, 41*(4), 226–232. https://doi.org/10.1207/s15430421tip4104_4

Meyer, A., Gordon, D., & Rose, D. H. (2014). *Universal Design for Learning: Theory and Practice.* CAST Professional Publishing.

Mullis, I. V. S., Martin, M. O., Foy, P., Kelly, D. L., & Fishbein, B. (2020). *TIMSS 2019 International results in mathematics and science.* Boston College, TIMSS & PIRLS International Study Centre: https://timssandpirls.bc.edu/timss2019/international-results/

National Council of Teachers of Mathematics (NCTM). (2000). *Principles and Standards for School Mathematics.* NCTM.
https://hdl.handle.net/20.500.12365/17719

National Governors Association Centre for Best Practices (NGACBP) & Council of Chief State School Officers (CCSSO). (2010). *Common Core State Standards for Mathematics*. NGACBP & CCSSO Authors.

Niss, M. and Højgaard, T. (2019). Mathematical competencies revisited. *Educational Studies in Mathematics, 102*, 9-28. https://doi.org/10.1007/s10649-019-09903-9

Organisation for Economic Co-operation and Development (OECD). (2017). *PISA Assessment and Analysis Framework for Development: Reading, Mathematics and Science.* OECD Publishing.

Piggott, J. (2011). *Mathematics enrichment: What is it and who is it for? NRICH - Millennium Mathematics Project.* Cambridge University. https://nrich.maths.org/5737

Pólya, G. (1945). *How to solve it*. Princeton University Press.

United Nations Development Programme (UNDP). (2020). *Human Development Report 2020.* The Next Frontier: Human Development and the Anthropocene. UNDP.

Purwadi, I., Sudiarta, I., & Suparta, I. N. (2019). The Effect of Concrete-Pictorial-Abstract Strategy toward Students' Mathematical Conceptual Understanding and Mathematical Representation on Fractions. *International Journal of Instruction, 12*(1), 1113-1126. https://doi.org/10.29333/iji.2019.12171a

Santos-Trigo, M. (2024). Problem solving in mathematics education: tracing its foundations and current research-practice trends. *ZDM–Mathematics Education, 56*(2), 211–222. https://doi.org/10.1007/s11858-024-01578-8

Schleicher, A. (2019). *PISA 2018: Insights and Interpretations*. OECD Publishing.

Schoenfeld, A. H. (2014). *Mathematical Problem Solving*. Elsevier.

Shuxratovna, R. N. (2024). Pedagogical possibilities of implementing the CPA (Concrete-Pictorial-Abstract) approach. *International Journal of Pedagogics, 4*(2), 68–76. https://doi.org/10.37547/ijp/volume04issue02-13

Silver, E. A., Kilpatrick, J., & Schlesinger, B. (1990). *Thinking through mathematics*. College Board Publications.

Skemp, R. R. (1976). Relational Understanding and Instrumental Understanding. *Mathematics Teaching, 77*, 20–26. https://ci.nii.ac.jp/naid/10010963202

Slavin, R. E., Lake, C., Davis, S., & Madden, N. A. (2011). Effective programmes for struggling readers: A best-evidence synthesis. *Educational Research Review*, *6*(1), 1–26. https://doi.org/10.1016/j.edurev.2010.07.002

Solomon, T., Dupuis, A., O'Hara, A., Hockenberry, M.-N., Lam, J., Goco, G., Ferguson, B., & Tannock, R. (2019). A cluster-randomised controlled trial of the effectiveness of the JUMP Math programme of maths instruction for improving elementary maths achievement. *PLoS ONE, 14(*10), e0223049. https://doi.org/10.1371/journal.pone.0223049

Stanton, E. A. (2007). The Human Development Index: A History. *Political Economy Research Institute (PERI) Working Papers Series No. 127.* University of Massachusetts Amherst. https://doi.org/10.7275/1282621

Stephan, M. and Akyuz, D. (2022). Semiotics from a social constructivist perspective. *International Journal of Science and Mathematics Education*, *20*, 1499–1519. https://doi.org/10.1007/s10763-021-10212-y

Stigler, J. W. and Hiebert, J. (1999). *The Teaching Gap: Best Ideas from the World's Teachers for Improving Education in the Classroom*. The Free Press.

Suárez-Pellicioni, M., Núñez-Peña, M. I., & Colomé, À. (2016). Math anxiety: A review of its cognitive consequences, psychophysiological correlates, and brain bases. *Cognitive, Affective, & Behavioural Neuroscience, 16*(1), 3-22. https://doi.org/10.3758/s13415-015-0370-7

Tashtoush, M. A., Wardat, Y., Aloufi, F., & Taani, O. (2022). The effect of a training programme based on TIMSS on developing the levels of habits of mind and mathematical reasoning skills among pre-service mathematics teachers. *Eurasia Journal of Mathematics, Science and Technology Education, 18*(11), em2182. https://doi.org/10.29333/ejmste/12557

Teig, N., Scherer, R., & Olsen, R. V. (2022). A systematic review of studies investigating science teaching and learning: over two decades of TIMSS and PISA. *International Journal of Science Education, 44*(12), 2035-2058. https://doi.org/10.1080/09500693.2022.2109075

Thiyagu, K. (2013). Effectiveness of Singapore math strategies in learning mathematics among fourth standard students. *Vetric Education, 1*, 1-14.

Vilalta, A. (2021). A project to develop mathematical competence in the primary school classroom*. Uno: Revista de Didáctica de las Matemáticas, 92*, 73-79.

Valencia Álvarez, A. B. and Valenzuela González, J. R. (2017). What types of mathematical problems are calculus students exposed to? An analysis of textbooks. *Mathematics Education, 29*(3), 51-78.https://doi.org/10.24844/em2903.02

Welch, B. L. (1947). The generalisation of 'Student's' problem when several different population variances are involved. *Biometrika*, *34*(1–2), 28–35. https://doi.org/10.1093/biomet/34.1-2.28

Date received: 28 September 2024
Date of review: 8 October, 2024
Date of acceptance: 14 March 2025